

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 14:58:33 ; Search time 99 seconds  
(without alignments)  
912.307 Million cell updates/sec

Title: US-09-745-506-37  
Perfect score: 1799  
Sequence: 1 MDKALSLINDFASLSFAE.....LENKINIILSETDRDPLQVY 350

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1316	73.2	340	11	08C608	08C608	mus musculus
2	707	39.3	159	11	09D146	09D146	mus musculus
3	358.5	19.9	376	16	08R972	08R972	thermoanaer
4	312	17.3	366	16	08CSD9	08CSD9	staphylococ
5	309.5	17.2	372	16	08EP20	08EP20	oceanobacti
6	279	15.5	366	16	08NMB9	08NMB9	staphylococ
7	271.5	15.1	258	16	08RE15	08RE15	fusobacteri
8	256	14.2	242	5	08SV52	08SV52	encephalito
9	245	13.6	433	16	08FNK9	08FNK9	corynebacte
10	230	12.8	380	16	08DNV4	08DNV4	streptococc
11	203.5	11.3	309	16	08G5M5	08G5M5	bifidobacte
12	189.5	10.3	262	16	08DT99	08DT99	streptococc
13	185.5	10.1	265	2	09LAG2	09LAG2	lactobacill
14	181.5	9.7	262	16	08DZA7	08DZA7	streptococc

17	173	9.6	262	16	08E4W8	08E4W8	streptococc
18	151.5	8.4	298	17	08RM42	08RM42	methanocarc
19	143.5	8.0	253	16	08RUT6	08RUT6	mycoplasma
20	138	7.7	306	17	08PTJ7	08PTJ7	methanocarc
21	135.5	7.5	248	17	08RW44	08RW44	methanopyru
22	114	6.3	250	16	08EDX0	08EDX0	shewanella
23	114	6.3	1260	5	09XZC5	09XZC5	cryptospori
24	110.5	6.1	251	2	08KRX3	08KRX3	vibrio fusc
25	109	6.1	249	16	08D2D0	08D2D0	wiggleswort
26	105.5	5.9	252	17	08U1Y9	08U1Y9	pyrococcus
27	105.5	5.8	566	16	09WZU1	09WZU1	thermocoga
28	104.5	5.8	456	16	08ECAB	08ECAB	shewanella
29	101.5	5.6	850	16	08DLR3	08DLR3	synecococc
30	100	5.6	221	1	050523	050523	methanobact
31	99	5.5	564	11	08R2U2	08R2U2	mus musculu
32	98.5	5.5	745	11	08CGH7	08CGH7	mus musculu
33	98.5	5.5	1036	2	08GRU1	08GRU1	actinobact
34	98.5	5.5	1199	11	054774	054774	mus musculu
35	98	5.4	1093	16	09K6B3	09K6B3	bacillus ha
36	97.5	5.4	811	16	08BHP1	08BHP1	xanthomonas
37	96.5	5.4	559	16	08XJY8	08XJY8	clostridium
38	96	5.3	238	17	P94953	P94953	methanopyru
39	96	5.3	233	17	08T2C0	08T2C0	methanopyru
40	96	5.3	261	16	08DKT4	08DKT4	synecococc
41	96	5.3	597	5	09VV09	09VV09	drosophila
42	96	5.3	1286	4	096KN7	096KN7	homo sapien
43	96	5.3	1286	4	096Q87	096Q87	homo sapien
44	95	5.3	452	17	08TGZ4	08TGZ4	methanopyru
45	94.5	5.3	325	16	08XJ57	08XJ57	clostridium

#### ALIGNMENTS

RESULT 1	ID	08C608	PRELIMINARY;	PRT;	340 AA.
AC	08C608	08C608			
DT	01-MAR-2003 (TREMBLrel. 23, Created)				
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)				
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)				
DE	N991 Interacting factor 3-like 1.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	SPRAIN-C57BL/6J; TISSUE=Testis;				
RX	MEDLINE=22354683; PubMed=12466851;				
RA	The FANTOM Consortium.				
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;				
RT	"Analysis of the mouse transcriptome based on functional annotation of				
RT	60,770 full-length cDNAs."				
RL	Nature 420:563-573(2002).				
DR	EMBL: AK076788; BAC36481.1;				
SQ	SEQUENCE 340 AA; 37822 MW; 82DC63DB90B684E CRC64;				
Query Match	73.2%; Score 1316; DB 11; Length 340;				
Best local Similarity	84.6%; Pred. No. 3.6e-100;				
Matches 253; Conservative 21; Mismatches 25; Indels 0; Gaps 0;					
OY	1 MDKALSLINDFASLSFAESMDNGLVEPSPHVTNLTFTNDLTEEYMEYVLRKAD 60				
DB					
OY	61 LIISYHPPIFRPKRTITWNTWKERIVIRALENVGIVSPPTAYDAAPQGNMVLAKGIGA 120				
DB					
OY	87 FILSYHPPIFRPKRTITWNTWKERIVIRALENVGIVSPPTAYDAAPQGNMVLAKGIGA 146				
OY	121 CTSRPHPSPKAPYPTGEGNHRVEFNVYTDLDKNSAVKIGDGVSTSPSARTGNEQDT 180				
DB					
OY	147 CTRPPIHPSPAPYPTGEGNHRVEFNVYTDLDKNSAVKIGDGVSTSPSARTGNEQDT 206				
DB					



RC STRAIN-ATCC 12228;  
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,  
RA Chen Z., Men Y.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AE016748; AA004845.1; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 366 AA; 41639 MW; BF73A88EE77462A1 CRC64;

Query Match 17.3%; Score 312; DB 16; Length 366;  
Best Local Similarity 24.9%; Pred. No. 2.3e-17;  
Matches 100; Conservative 80; Mismatches 130; Indels 92; Gaps 16;

QY 1 MDKALLSLNDFASTFAESMDNVGLVSPSPHTVTLTNDLDEEVEYLQKAD 60  
DB 1 MKISELMEVLNNHVPVPHQAESMDNVGLI-GNDKLDITGLITLTDCTDDVNNALNTN 59  
QY 61 LILSYHPIPRPKRITWTWKERLVIRALENVGIYSPTAYDAAPQGVNMLAKGLA 120  
DB 60 TIIAHHPILFKGVKRIYEDGYGS-ITRKLQNNINILALHTNLDVNPCKGVIRMLADQIGL 118  
QY 121 CTSRPIHPKAPRYPT-----EGNHRVEENV----- 147  
DB 119 ENISMINTSSYYKYQVETIPKNYIEDFKDSLNEGLAEGNYEGCFEESKGGQRPVG 178  
QY 148 ---YTQDLKVASAVKIDGVSVTSFSAATGNEQTRINLCTOKALMO-----VVD 196  
DB 179 DASPIYCKLD---SIEYDEIKL-EFMK-GNE-----LETKRAILDNHPEYTPVD 226  
QY 197 FLSRNKQLYQTEILSLERPLLHTGMGLCTLDESVSATMIDRIKRLILSHIRLALG 256  
DB 227 FIKMKE-----SEYGLIGLQNLQMTLDEFSEYAKQLNIPSVRY--- 268  
QY 257 VGRTEISOYKYNALCAGSSGVAGVE-----ADLYLGEMSHDITLDAAGQINVL 309  
DB 269 TGGH-DSPIKKVAIIIGSGSI---GFEYKASQGLADVPYVGIKHDALDAKIQNVNLLD 323  
QY 310 CEHSNTERGFLSDRLMDLSHL---ENKNIILSETDRDPLQ 348  
DB 324 INH-YSEYVKEGKLKELKELFKYENQPIYASEINTDPFK 364

RESULT 5  
Q8EPZ0 PRELIMINARY; PRT; 372 AA.  
AC O8EPZ0:  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Hypothetical conserved protein.  
GN O81940.  
OS Oceanobacillus thelyensis.  
OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.  
ON NCBL\_TaxID=182710;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-HTE831 / DSM 14371 / JCM 11309;  
RX MEDLINE=2220767; PubMed=12235376;  
RA Takami H., Takaki Y., Uchiyama T.;  
RT "Genome sequence of Oceanobacillus thelyensis isolated from the Iheya  
RT Ridge and its unexpected adaptive capabilities to extreme  
RT environments.";  
RL Nucleic Acids Res. 30:3927-3935(2002).  
DR EMBL: AP004599; BAC13896.1; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 372 AA; 41869 MW; B66B9C99C8749C43 CRC64;

Query Match 17.2%; Score 309.5; DB 16; Length 372;  
Best Local Similarity 28.1%; Pred. No. 3.8e-17;  
Matches 105; Conservative 70; Mismatches 152; Indels 47; Gaps 13;

QY 9 SLNDFSLSFASMDNVGLVSPSPHTVTLTNDLDEEVEYLQKADLILSYHPP 68  
DB 14 AMDVAPKHLAYDMDNVGLQIGSFNAITSVL-ITLDVTSVVDALIKGMNIIIAHSL 72

QY 69 IFRPMKRTWTWKERLVIRALENVGIYSPTAYDAAPQGVNMLAKGLACTSRPIHP 128  
DB 73 LFQPLKQINNDTKGRVQKRLKNDITVYASHRLDANNGVMDMLATELKLHTTTP-- 130  
QY 129 SKAPNPTTEGNHREVEENVYTTOD-LDKVASAVKIDGVSVTSFSART----- 174  
DB 131 --VPVY-NEKNYV---YIVPESHIEKINQAFASGAGYIGNVSHOTFSPQGGTFKPL 184  
QY 175 -----GNE-----EQTINLNCOKALMOY-DELSRNKQLYQTEILSLERPLLLH 220  
DB 185 EGTDPYIGEQDKITTYDEYKIEITYPEISLNVKSTLSAHPEEPAYDIYPLENK-TTN 243  
QY 221 TGMGRCTLDESVSATMIDRIKRLILSHIRLALGVGRLESQVKVALCAGSSVLQ 280  
DB 244 FGIGRIGNLEPTPLDSFITYKTFQPLEIR---ISGNKDKIKRVALIGSGGEYVS 299  
QY 281 GV---EADLYLGEMSHDITLDAAGQINVLCEHSNTERGFLSDRLMDLSHL-NKIN 336  
DB 300 HAMRKADVYITGDMFPHIQDAENGELTYIDAGH-YIEKVMYTYTKQLAQTMELDKDF 358  
QY 337 IILSETDRDPLQYV 350  
DB 359 IEVSEINTDPEQYV 372

RESULT 6  
Q8NMB9 PRELIMINARY; PRT; 366 AA.  
AC Q8NMB9;  
DT 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Hypothetical protein MM1511.  
GN MM1511.  
OS Staphylococcus aureus (strain MM2).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
ON NCBL\_TaxID=196620;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22040717; PubMed=12044378;  
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,  
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,  
RA Yamamoto K., Hiramoto K.;  
RT "Genome and virulence determinants of high virulence community-  
RT acquired MRSA.";  
RL Lancet 359:1819-1827(2002).  
DR EMBL: AP004827; BAB95376.1; -  
DR InterPro: IPR002678; DUF34.  
DR Pfam: PF01784; DUF34.1.  
DR TIGRfams: TIGR00486; TIGR00486.1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 366 AA; 41153 MW; FALF28762D878C79 CRC64;

Query Match 15.5%; Score 279; DB 16; Length 366;  
Best Local Similarity 25.9%; Pred. No. 1.2e-14;  
Matches 102; Conservative 69; Mismatches 151; Indels 72; Gaps 16;

QY 1 MDKALLSLNDFASTFAESMDNVGLVSPSPHTVTLTNDLDEEVEYLQKAD 60  
DB 1 MKIADLMTLTDHHPFSTAESMDNVGLIIDEDEVYGVLTALDCTLEVNNEALIEKGYN 59  
QY 61 LILSYHPIPRPKRITWTWKERLVIRAL-ENRVGIYSPTAYDAAPQGVNMLAKGLG 119  
DB 60 TIIASHHPILFKGVTSKLANGY--GLITRKLQHDINILAHNTNLDVNPYGVNMLAKAWG 117  
QY 120 ACTSRPIHPKAPRYPT-----EGNHRVEF-----NVNYTQDL 152  
DB 118 LKNISITNNQDVIYKQVETIPKDNVGFKDKLSENGLAEGNVEYCFEESKGGQRPV 177  
QY 153 DKVASAVKIDGVSVTSFSAATGNEQTRINLCTOKALMOYVDFLSRNKQLYQK----- 207  
DB 178 GEANPTIGQIDKIEYV-----DEVKIEFMIDAYQK-----SRADQLIKOYHPYE 221

OY 208 ---TEILSLERPLLIHTGMRICLTDESVALTMIDRIKRLKLSHRLALGVTLESQ 264  
 Db 222 TPVDFEIRIKOTSILY--GLGVMAEVDNQMTLEDFADIAKSKINIPSVRF---VGES-NQK 275  
 OY 265 VKVVALCAGSG-----SSVLOGVENDLYLTGEMSHDPLDAASOGINVLCEHSNTENG 319  
 Db 276 IKRIALIGSGSIGEYQAVQOG--ADVFTGDIKIHDAIDAKIHGVNLIDINH-VSEYVM 332  
 OY 320 LSDRLMDLSHLE-NKINIL--SETDRDPLQVY 350  
 Db 333 KEGKLTLMNRFNPFKINIDVASTINDPQYI 366

RESULT 7

O8RE15 PRELIMINARY; PRT; 258 AA.  
 AC O8RE15;  
 DT 01-JUN-2002 (Tremblrel. 21, Created)  
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE NGG1-Interacting factor 3.  
 GN FNJ316.  
 OS Fusobacterium nucleatum (subsp. nucleatum).  
 OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;  
 CC Fusobacterium  
 OX NCBI\_TaxId=76856;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 25586;  
 RX MEDLINE-21886394; PubMed-11889109.  
 RA Kapriel V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,  
 RA Bhattacharya A., Bartman A., Gardner W., Gschick G., Zhu L.,  
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,  
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,  
 RA Fosteijn M., Kyrides N., Overbeek R.;  
 RA "Genome sequence and analysis of the oral bacterium Fusobacterium  
 RT nucleatum strain ATCC 25586";  
 RL J. Bacteriol. 184:2005-2018(2002).  
 DR EMBL; AE010636; AAL95512.1; -  
 DR InterPro; IPR002678; DUF34.  
 DR Pfam; PF01784; DUF34; 1.  
 DR TIGRfams; TIGR00486; TIGR00486; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 258 AA; 30127 MW; 4A6A2B53B64F4280 CRC64;

Query Match 15.1%; Score 271.5; DB 16; Length 258;  
 Best Local Similarity 23.3%; Pred. No. 3e-14;  
 Matches 74; Conservative 61; Mismatches 88; Indels 95; Gaps 7;

OY 19 AESMDNGLVEPSPHTVNTLELTNDLVEEVEVLQKADLLISYHPPIFRPKRITW 78  
 Db 19 AEEWNIIGLLIGDYDE-VKKIQFSLDITLESIEKVDMLTTHPIFKAIKDYTE 77  
 OY 79 NTKWKRLVIRALENVGIYSPHTADAPQGVNNMLAGLGACTSRPIHPSKAPYPIEG 138  
 Db 78 QNIIKAKRLKLNKINIVSIHTNLDSSIEGLNDVLRKIG----- 118  
 OY 139 NHRVEFNVTQDDLDKMSAVAGIDGVSPTSARTGNEQTRINLCTOKALMQVDFL 198  
 Db 119 -----ISEKIIIDFBEK-----NC----- 133  
 OY 199 SRNKQLYKTEILSLERPLLIHTGMRICLTDESVALTMIDRIKRLKLSHRLALGV 258  
 Db 134 -----GIGRIFKLNEEKNLKFIEELKLIILRLV---IS 167  
 OY 259 RTLESQVVAALCAGSSSVLQV---EADLYLTGEMSHDPLDAASOGINVLCEHSNT 315  
 Db 168 NDLNKKIKKVALINSANMYMKAKKKEKVDLFTIDGVSYHDLALENGSLVIGDHES 227  
 OY 316 ERGFLSDL-RDMLDSHLE 332  
 Db 228 EFFYEILIEELKDNLE 245

RESULT 8

O8SV52 PRELIMINARY; PRT; 242 AA.  
 AC O8SV52;  
 DT 01-JUN-2002 (Tremblrel. 21, Created)  
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
 DE Hypothetical protein ECU07\_0100.  
 GN ECU07\_0100.  
 OS Encephalitozoon cuniculi.  
 OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.  
 OX NCBI\_TaxId=6035;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GB-M1;  
 RA Genoscope;  
 RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GB-M1;  
 RX MEDLINE-21576510; PubMed-11719806;  
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarar F.,  
 RA Prensier G., Barbe V., Peyretailade E., Brotlier P., Wincker P.,  
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,  
 RA Weissenbach J., Vivas C.P.;  
 RA "Genome sequence and gene compaction of the eukaryote parasite  
 RT Encephalitozoon cuniculi";  
 RL Nature 414:450-453(2001).  
 DR EMBL; AL590447; CAD25542.1; -  
 DR InterPro; IPR002678; DUF34.  
 DR Pfam; PF01784; DUF34; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 242 AA; 26877 MW; D95448A74B97B7D6 CRC64;

Query Match 14.2%; Score 256; DB 5; Length 242;  
 Best Local Similarity 24.1%; Pred. No. 5.2e-13;  
 Matches 85; Conservative 56; Mismatches 98; Indels 114; Gaps 13;

OY 1 MDKLALSLNDFASLSPAES-WDNVGLVEPSPHTVNTLELTNDLVEEVEVLQKKA 59  
 Db 1 MSIHEVSAIDRFASLRAKTDMDVGVYVDSGTD--KKILITIDLEPVLEECIHGV 58  
 OY 60 DLISYHPPIFRPKRITWNTWKRLVIRALENVGIYSPHTADAPQGVNNMLAKGIG 119  
 Db 59 KNVVAYHPIFKAVKVL--GSKESIVIGICINIVFTPSALDPL---MNTYV---YN 109  
 OY 120 ACTSRPIHPSK--APNYPLEGNHREFNVTQDDLDKMSAVAGIDGVSPTSARTGNE 177  
 Db 110 MINSGPFFYKKKNGGPTSTIGN-----AIRLERSGLE 143  
 OY 178 EQTRINLCTOKALMQVDFLSRNKQLYKTEILSLERPLLIHTGMRICLTDESVALT 237  
 Db 144 S---PRICLARHTMESV-----PERNHVVG-----AT 169  
 OY 238 MIDRIKRLKLSHRLALGVGTLESQVYVVALCAGSSSVLQGVENDLYLTGEMSHD 297  
 Db 170 F-----RNVLLKN-----SIVLTGEMSHDL 190  
 OY 238 LDAASOGINVLCEHSNTERGFLSDLRDLMDLSHLENKINIISETDRDPLQVY 350  
 Db 191 LSCIANGASVILMEHSNSERICLGHISAKLEELP-EYEIVSKNDKDPVIL 242  
 RESULT 9  
 O8FNK9 PRELIMINARY; PRT; 433 AA.  
 AC O8FNK9;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Conserved hypothetical protein.



Db 6 VIQAYEAFQPEFSMEDSGGLQI-GTLDKIGQVWALDIREVTALEKGVDLTIK 64  
QY 66 HPIIFRPMKRITVNTWKERLVIRALENRVGIYSPHTAYDAPOGVNMLAKGACTSRP 125  
Db 65 HAPIFRPKIDLLASRPQNIYIDLIKHDIAVYVSHNIDIVENGNDWFCOML----- 117  
QY 126 IHPKSKAPNTEEGNHRREFVNVNTQDLDKYSMAVKGIDGVSVSFSARTGNEQTRINLN 185  
Db 118 -----GIEETTYLOETGPER----- 132  
QY 186 CTOKALMQVVDFTLSRNKQLYOKTEILSEKPLIHTGMRLCTIDESVSATMIDRIKRH 245  
Db 133 -----GIGHIGNI-QPQTFWELAQYKQY 155  
QY 246 LKISHIRLALGVGTLESQYKVALCAGSSSVLQGV---EADLYLTGEMSHDTPDAAS 302  
Db 156 FDLDSLHMYHQENDLQKPSRAVICGSGSEFKDALAKGADVYITGDYIYHTAQMLLS 215  
QY 303 OGIVNICEHSNTERGFLSDRLDMLDHLNKK---INIIISERTDRLDQ 348  
Db 216 DGLALDPGH-YIEVLFVERKIALLTQWKKEKWELEILDSQASTNPF 263

## RESULT 12

08G5M5 PRELIMINARY; PRT; 309 AA.  
AC 08G5M5;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Conserved hypothetical protein with duf34.  
GN BL0984.  
OS Bifidobacterium longum.  
OC Bacteria; Actinobacteria; Actinobacteriales; Bifidobacteriales;  
OC Bifidobacteriaceae; Bifidobacterium.  
OX NCBI\_Taxid=216816;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCC 2705;  
RX MEDLINE=22294977; PubMed=12381787;  
RA Schell M.A., Karamitanzou M., Snel B., Vilanova D., Berger B.,  
RA Pessi G., Zwaalen M.-C., Desiere F., Bork P., Delley M.,  
RA Pridmore R.D., Arigoni F.;  
RT "The genome sequence of Bifidobacterium longum reflects its adaptation  
RT to the human gastrointestinal tract."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).  
DR EMBL; AE014721; AAN24792.1;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 309 AA; 33851 MW; 68344899059F0F03 CRC64;

Query Match 11.3%; Score 203.5; DB 16; Length 309;  
Best Local Similarity 21.8%; Pred. No. 1.6e-08;  
Matches 67; Conservative 45; Mismatches 106; Indels 89; Gaps 5;

QY 2 DLKALLSLNDFASLFAESMDVNGLLVEPSRPHVTNLTFLNLTLEEMVEVLQKKDL 61  
Db 6 NLKQVIVLETLFLYLRKAEEMDEPGLY-GDLSHDVHRTIYAAOPTSAITDKAATGADL 64  
QY 62 ILSTYHPIFRPMKRITVNTWKERLVIRALENRVGIYSPHTAYDAPOGVNMLAKGAC 121  
Db 65 LITHHPLFEFRVHETSGIFRGDIYRLYHGCGLWGHGTNADAAYRGVGAADYFELI 124  
QY 122 TSRIHPKSKAPNTEEGNHRREFVNVNTQDLDKYSMAVKGIDGVSVSFSARTGNEQTR 181  
Db 125 DQKPLVLPIDAN---ASHPV----- 141  
QY 182 INLNCQKALMQVVDFTLSRNKQLYOKTEILSEKPLIHTGMRLCTIDESVSATMIDR 241  
Db 142 -----GLGRVGRLPKPVALKDFQR 161  
QY 242 IKRHLKISHIRLALGVGTLESQ--VKVVALCAGSSSVLQGV---DLYLTGEMSHD 296  
Db 162 VFDEYSDHGHTSALGIGVCGSDTFVQYVALIPSGSDLEFNEVRAITGVYVTSDLRHP 221

QY 297 TLDASQ 303  
Db 222 VTDALIQ 228

## RESULT 13

08DTP9 PRELIMINARY; PRT; 262 AA.  
AC 08DTP9;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Conserved hypothetical protein.  
GN SMU.1463C.  
OS Streptococcus mutans.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_Taxid=1309;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=UA159 / ATCC 700610 / Serotype C;  
RX MEDLINE=22295063; PubMed=12397186;  
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,  
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Ian S., Qian Y.,  
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;  
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental  
RT pathogen."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).  
DR EMBL; AE014978; AAN59121.1;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 262 AA; 29740 MW; A5740C558BEF248B CRC64;

Query Match 10.5%; Score 189.5; DB 16; Length 262;  
Best Local Similarity 20.8%; Pred. No. 1.8e-07;  
Matches 66; Conservative 50; Mismatches 100; Indels 101; Gaps 10;

QY 37 VNTFLNLTLEEMVEVLQKKADLLISYHPIFRPMKRITVNTWKERLVIRALENRVGI 96  
Db 36 IKRIMALDVETTVAEIEKRVLLYKAPITRPLANTLY-ETANQHHIIFNLKHDIAV 94  
QY 97 YSPHTAYDAPOGVNMLAKGAGCTSRPIHPSKAPNTEEGNHRREFVNVNTQDLDKYM 156  
Db 95 YVSHNTNIDYVDGLDMFCDDLIDKNNRIILSPSD----- 129  
QY 157 SAVKIGDVSYSFSARTGNEQTRINLNCQKALMQVVDFTLSRNKQLYOKTEILSEK 216  
Db 130 -----DY----- 131  
QY 217 LLHTGMRLCTIDESVSATMIDRIKRHLKLSHRL-ALGVGTLESQYKVALCAGSG 275  
Db 132 -----GIGRVGDI-SPUSFEDLAKVKKIFMLDSVRLVSGEENPLISR--VAICGSG 182  
QY 276 SSVLQGV---EADLYLTGEMSHDTPDAASOGIVNICEHSNTERGFLSDRLMD--SH 330  
Db 183 OSFYQALTKGAQYVITGDYIYHTAQMLTNGLLALDPGH-HIVLFLVRELAERFQWTSC 241  
QY 331 LEN-KINIIISERTDTP 346  
Db 242 QENNDITLESQVNTNP 258

## RESULT 14

P95777 PRELIMINARY; PRT; 262 AA.  
AC P95777;  
DT 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE ORF4 protein.  
OS Streptococcus mutans.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.



**THIS PAGE BLANK (USPTO)**